Expanding Acceptable Transfer Requirements:
Transfer Instructions for Permanent Electronic Records

RECORDS IN PORTABLE DOCUMENT FORMAT (PDF)

1.0   PREFACE

As part of the National Archives and Records Administration's (NARA) electronic government (E-Gov) initiative, and in cooperation with other Federal agencies, NARA is issuing guidance to supplement current requirements in 36 CFR 1228.270 for transferring permanent electronic records to NARA. This guidance expands currently acceptable formats to enable the transfer of records in Portable Document Format (PDF) to NARA.

This guidance provides transfer requirements for:

- all records in PDF (sections 3.1 and 3.2),
- records converted to PDF from their native electronic formats (e.g., office automation products) (section 3.3),
- records converted to PDF from scanned paper or image formats such as TIFF (section 3.4).

Additionally, sections 3.5 and 4.0 provide transfer requirements including transfer documentation and related information on how to transfer PDF records to NARA.

1.1   Effective Dates

This guidance applies to all PDF records that have been appraised or scheduled for permanent retention. The effective dates are based on when the PDF records are created.

A.   *March 31, 2003*. The requirements in this guidance are effective March 31, 2003, for <u>all</u> permanent PDF records created prior to April 1, 2004.

B.   *April 1, 2004*. The effective date of the following <u>additional</u> provisions has been deferred until April 1, 2004, to allow agencies time to implement them:

- deactivate all security settings before transfer (section 3.2.1.2),
- embed all referenced fonts (section 3.3.1.2).

NARA recognizes that legacy records and records whose disposition is changed from temporary to permanent may present unique circumstances for agencies. Any agency that has permanent PDF records that do not meet the requirements in this guidance should contact the NARA appraisal archivist assigned to that agency to determine the most appropriate medium and format for transfer (see section 6.0).

2.0   BACKGROUND

PDF is a priority electronic records format identified by NARA and partner agencies as part of the Electronic Records Management (ERM) initiative, one of the twenty-four E-Gov initiatives under the President's Management Agenda. A major goal of this initiative is to provide the tools for agencies to access electronic records for as long as required and to transfer permanent

electronic records to NARA for preservation and future use by government and citizens. NARA has previously issued transfer guidance for email with attachments and scanned images of textual records under this ERM initiative.

These guidance documents and additional information about the ERM E-Gov initiative can be found on NARA's web site.[1]

3.0   Transfer Requirements for PDF Records

NARA will accept transfers of PDF records that have been scheduled as permanent records on a *SF 115, Request for Records Disposition Authority.* Any agency that has permanent PDF records that do not meet the requirements in this guidance, should contact the NARA appraisal archivist assigned to that agency (see section 6.0).

To facilitate preservation processing and future access to these records, agencies must comply with the following minimum requirements:

3.1   PDF File Specification for All PDF Records

3.1.1   PDF records must comply with PDF versions 1.0 through 1.4 (i.e., all existing PDF versions as of the effective date of this guidance) and meet the requirements outlined in sections 3.2 through 3.4.

3.1.2   NARA periodically will update the list of acceptable PDF versions as required.

3.2   General Requirements for All PDF records

3.2.1   Security Requirements

3.2.1.1 PDF records must **not** contain security settings (e.g., self-sign security, user passwords, and/or permissions) that prevent NARA from opening, viewing or printing the record.

3.2.1.2 In addition, PDF records created after April 1, 2004, must have all security settings deactivated (e.g., encryption, master passwords, and/or permissions) prior to transfer to NARA. Deactivating security settings ensures NARA's ability to support long term migration and preservation of the records.

---

[1] http://www.archives.gov /records_management/ initiatives/erm_overview.html.

3.2.2   Review of Special Features

Because of the complexities associated with certain PDF features, NARA will review PDF records containing special features on a case-by-case basis when the records are scheduled. Examples of special features include but are not limited to: digital signatures; links to other documents, files or sites; embedded files (including multimedia objects); form data; comments and/or annotations.

3.3   Requirements for Records Converted to PDF from Their Native Electronic Formats (e.g., office automation products)

3.3.1   Electronic records that have been converted to PDF from their native electronic formats must include embedded fonts to guarantee the visual reproduction of all text as created. All fonts embedded in PDF records must be publicly identified as legally embeddable (i.e., font license permits embedding) in a file for unlimited, universal viewing and printing.

3.3.1.1 PDF records that reference fonts other than the "base 14 fonts"[2] must have those fonts referenced in the record (i.e., as a minimum, subsets of all referenced fonts) embedded within the PDF file.

3.3.1.2 PDF records created after April 1, 2004, must have all fonts referenced in the record, including the "base 14 fonts," embedded within the PDF file. This requirement is met by having, as a minimum, subsets of all referenced fonts embedded within the PDF file.

3.4   Requirements for Scanned Paper or Image Formats Converted to PDF

3.4.1   Scanned images of textual paper records converted to PDF must adhere to the requirements in NWM 02.2003, MEMORANDUM TO AGENCY RECORDS OFFICERS: *Expanding Acceptable Transfer Formats: Transfer Instructions for Scanned Images of Textual Records* (Scanned Images Transfer Guidance)*,* dated December 23, 2002.[3]

Any agency that has PDF records that have not been scanned according to the minimum image quality specifications in the NWM 02.2003 guidance, should contact the NARA appraisal archivist assigned to that agency (see section 6.0).

3.4.2   PDF records that contain embedded searchable text based on Optical Character Recognition (OCR) must be identical in content and appearance to the source document. NARA understands that the ability to embed OCR'd text in PDF records enhances access to the records. While NARA will accept PDF records

---

[2] The base 14 fonts are: Courier (Regular, Bold, Italic, and Bold Italic), Arial MT (Regular, Bold, Oblique, and Bold Oblique), Times New Roman PS MT (Roman, Bold, Italic, and Bold Italic), Symbol, and ZapfDingbats.

[3] http://www.archives.gov/records_management/policy_and_guidance/nwm02_2003.html

with uncorrected OCR'd text, it will not accept PDF records resulting from OCR processes that either alter the content or degrade the quality of the original bit-mapped image.

3.4.2.1 NARA will accept PDF records that have been OCR'd using processes that do not alter the original bit-mapped image. An example of an output process that accomplishes this requirement is *Searchable Image – Exact*.

3.4.2.2 NARA will <u>not</u> accept PDF records that have been OCR'd using processes that substitute OCR'd text for the original scanned text within the bit-mapped image. Such OCR processes may involve loss of data through imprecise interpretation of scanned characters. Examples of output processes that use this prohibited technique include *Formatted Text and Graphics* and *PDF Normal*.

3.4.2.3 NARA will <u>not</u> accept PDF records that have been OCR'd using processes that use lossy compression to reduce file size (e.g., JPEG). Such OCR processes degrade the quality of the original image and may make such images unsuitable for archival preservation. An example of an output process that uses this lossy compression technique for color and grayscale images is *Searchable Image – Compact*.

3.5   Transfer Documentation

This guidance supplements transfer documentation requirements in 36 CFR 1228.270(e) to ensure that transfers of records in PDF are clearly identified and described. Agencies must also submit a signed *Standard Form 258, Agreement to Transfer Records to the National Archives of the United States (SF 258)*, as required by 36 CFR 1228.272.

3.5.1   For each transfer, agencies must supply documentation that identifies the software used to create the PDF records (if available) and the version(s), and the operating system (if available) and version(s).

3.5.2   Agencies must provide all external finding aids for the transferred PDF records (e.g., indexes; descriptive, administrative, or technical metadata; and/or databases of OCR'd text) in formats approved by NARA, with the appropriate documentation required by 36 CFR 1228.270(e).

3.5.3   When an agency has developed standards or guidelines to assist in formatting, validating, or accessing PDF records (including recommended software or quality settings, and/or guidelines for embedding metadata within PDF records), a copy of these standards or guidelines must be included with the transfer.

3.5.4   PDF records converted from scanned images also must adhere to the transfer documentation requirements in section 3.3 of the Scanned Images Transfer Guidance.

4.0   TRANSFER MECHANISMS

4.1   Agencies may transfer PDF records using any of the approved media or methods listed in 36 CFR 1228.270(c) that became effective January 29, 2003[4].

4.2   PDF records must not be compressed (e.g., Winzip, PKZIP) or aggregated (e.g., TAR) for purposes of transfer unless NARA has approved the transfer in compressed or aggregated form in advance. In such cases, NARA may require the agency to provide the software to decompress the records [see 36 CFR 1228.270(d)].

5.0   LEVELS OF ACCESS

NARA will provide access to the creating agency and to all researchers requesting PDF records accessioned from Federal agencies, subject to review of content for FOIA exemptions as is feasible. While compliance with these requirements will improve future access to accessioned PDF records, NARA's ability to provide access to certain records will vary according to their hardware and software dependencies. At the present time, NARA provides users with a copy of fully releasable electronic record files on any of the media currently approved by NARA. For PDF records transferred to NARA the user will be responsible for obtaining the necessary hardware and software to view the records.

6.0   CONTACT INFORMATION

For assistance in scheduling PDF records, or to discuss how to handle permanent PDF records that do not meet the specifications in section 3.0, contact your agency appraisal archivist in the Life Cycle Management Division (NWML). The NWML general telephone number is 301-837-3560.

For technical assistance in transferring PDF records to NARA, contact the Electronic and Special Media Records Services Division (NWME), 8601 Adelphi Road, College Park, MD 20740. The general telephone number is 301-837-3420.

---

[4] http://www.archives.gov/about_us/regulations/part_1228_l.html